LLNL-CONF-648477

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

WPEC SG38: Designing a New Format for Storing Nuclear Data

C. M. Mattoon, B. R. Beck

January 14, 2014

2014 ANS (American Nuclear Society) Annual Meeting
Reno, NV, United States
June 15, 2014 through June 19, 2014

WPEC SG38: Designing a New Format for Storing Nuclear Data

C. M. Mattoon and B. R. Beck

Lawrence Livermore National Laboratory: 7000 East Ave L-059, Livermore, CA, 94550.
mattoon1@llnl.gov

INTRODUCTION

The Evaluated Nuclear Data Format (ENDF) has been used for nearly 50 years to store and exchange evaluated nuclear data [1]. It provides a standard, computer-readable format where nuclear scientists and engineers can store data such as reaction cross-sections, outgoing energy and angle distributions for reaction products, thermal scattering parameters, and covariances. Data stored in ENDF serve as inputs for a variety of other codes, and can be processed into other formats for use in simulation codes such as MCNP [2]. ENDF therefore serves as a standard way for nuclear scientists and engineers to provide recommended nuclear reaction data for use in reactor design and other applications.

While ENDF has played an important role in nuclear physics and engineering for many years, it is showing signs of age. The format restricts both the types of data that can be stored, and the decimal precision that can be used to store them. ENDF was originally designed to be stored on 80-column punch cards, with 14 characters out of each line reserved for identifiers such as line numbers. Plain text documentation is allowed, but can only appear at the start of each file (that is, only once per projectile/target combination). ENDF files are difficult for humans to read, and frequently contain subtle errors that can remain hidden for years, partly due to the complexity of the format. Most importantly, a new generation of scientists and engineers are increasingly disinterested in learning the details of a legacy format like ENDF, preferring instead to work with modern software practices and tools (like XML) with which they are more familiar and that they find more useful. At a time when the nuclear data community is losing expertise to retirements, it needs to be willing to adopt new practices in order to attract a new generation of talent. For all these reasons, a modern replacement for ENDF has become an important priority.

Therefore, the Working Party on Evaluation Cooperation (or WPEC, a subsidiary of the OECD) has formed a new sub-group titled 'Beyond the ENDF format: a modern nuclear database structure' [3]. The goal of WPEC Subgroup 38 (SG38) is to oversee the design and implementation of a modern replacement for ENDF, that takes advantage of the many advances in computer technology since ENDF was created, and that is flexible, extensible, and both human- and computer-readable.

Simply designing the new format will not be sufficient. The widespread use of ENDF (including use in important applications like criticality safety and reactor licensing) means that changes to the format will have a significant impact. SG38 must therefore plan ahead and be prepared to help smooth the transition from ENDF to the new database format.

This paper presents the progress made so far by SG38. We begin by outlining the requirements that SG38 determined to be essential parts of an ENDF replacement, and then we present the progress that has been made towards meeting those requirements. In particular we will describe the Generalized Nuclear Data (GND) format [a] that is being developed at Lawrence Livermore National Laboratory (LLNL), which should serve as a good starting point for defining a new format.

REQUIREMENTS FOR THE NEW FORMAT

WPEC is an international organization that brings together nuclear data evaluators and users to improve the quality of evaluated nuclear data libraries. WPEC subgroups like SG38 are formed when a specific need is identified, and typically last a few years while that need is addressed. Subgroup 38 was formed in mid-2012 to work towards an ENDF replacement. The group (which is chaired by D. McNabb of LLNL) has met three times since its inception, twice at the Nuclear Energy Agency (NEA) in Issy-les-Moulineaux, France, and once at the Japan Atomic Energy Agency (JAEA) in Tokai, Japan.

The first task for SG38 was to define the requirements that the new format must meet. These requirements are meant to help capture the scope of the project: what types of data need to be handled by the new format, what tools are needed to enable easy user access to data in the new format, and any special problems that must be addressed during the transition to using the new format. The requirements identified by SG38 are:

- *Store the nuclear data in a hierarchy that reflects our understanding of nuclear reactions and decays, and that clearly and uniquely specifies all data*. This requirement does not mean that data will be stored in a particular computer language. Instead, it means that SG38 will define a general hierarchy that is suitable for storage in any hierarchical format (such as XML or HDF5 [4]). As much as possible, the hierarchy should be easy for anyone with a basic understanding of nuclear reactions to navigate and understand.
- *Support any projectile/target combination and any combination of reaction products (and subsequent*

decay products). While ENDF was originally designed to store only neutron-induced reactions, it has been expanded due to the evolving needs of users to include photonuclear and charged particle induced reactions. In the future other types of data (pion- and muon-induced reactions, for example) may be needed, so the new format must be able to extend to handle any particle interactions.

- *Support backwards-compatibility with ENDF for as long as possible*. Since ENDF is widely used around the world, SG38 must be careful to minimize the disruption to users caused by changing the format. Backwards-compatibility should help minimize this disruption, allowing users to continue using ENDF files for the near future. In practice, backwards-compatibility means two things: first, the new format must be able to handle all data types available in ENDF, and second, a translator for moving data automatically back and forth between ENDF and the new format is required. This translation tool will offer several benefits: nuclear data libraries can be provided in both formats with relative ease, new files can be translated into the old format for legacy codes, and code comparisons using the same data in two different formats should help find and fix bugs in codes and in data.
- *Include reusable low-level data containers that are general enough to be shared between data products*. While SG38 is focused on the evaluated nuclear reaction databases stored in ENDF, several other nuclear databases (such as the EXFOR database for storing experimental nuclear reaction data [5], and the ENSDF and RIPL databases for storing nuclear structure information like level schemes and decays [6,7]) are also in need of modernization. Well-designed basic data containers for storing data like x-y pairs, arrays, tables, parameterized curves, etc. should be made general enough to meet the needs of all these databases. Reusing these basic containers to store multiple types of data will help keep codes shorter, and hopefully easier to maintain.
- *Define APIs for reading and writing data in the new format.* An Application Programming Interface (API) should offer a standard way of accessing the data. User codes that utilize the API should be buffered against possible future changes in how data are stored, as long as the routines they use to access the data remain backwards compatible. If the data format does change, these users should not need to change to their codes unless they want to make use of new features.
- *Allow evaluators and data processors to record all the information needed to reproduce and extend their data*. This requirement is motivated by the fact that documentation in ENDF often does not contain sufficient information about how an evaluation was performed, what experimental data were used as inputs, and so on. By creating explicit containers for such documentation, we hope to improve the reproducibility of nuclear data evaluations in the future.
- *Encourage the elimination of inconsistent nuclear masses, levels and lifetimes by providing an external particle database and a way to link to it*. ENDF files contain redundant and frequently discrepant data. Most of this data relates to particle properties (masses in particular) that must be stored repeatedly throughout an ENDF file. In some cases, discrepancies cause problems such as non-conservation of energy. An independent particle properties format is being designed so that particle data can be stored as an external database. This particle database will help ensure consistency within an evaluation and between different evaluations. The reaction format should still allow individual evaluations to override parts of this central database when necessary.
- *Be governed by a long-term WPEC subgroup.* Since the new format (like ENDF) is intended to be used for worldwide exchange of nuclear data, it should be governed by an international body that oversees future changes and documentation. Since WPEC already serves as an international advisor for evaluators, it seems to be the best choice to host the format governance committee.

In addition to the requirements above, SG38 also advises that the new format should 1) *allow explicitly storing units and interpolation along with data*, 2) *support storing multiple representations of the same data simultaneously (e.g. both evaluated and processed data)*, and 3) *be forgiving, meaning that access routines for the new format must be able to recover gracefully and continue working if they encounter data containers not yet officially recognized as part of the format*.

PROGRESS TOWARDS ACCOMPLISHING SG38 GOALS

Nuclear data evaluators and users at the first SG38 meeting in December 2012 compiled the list of requirements above. This list has been written up in a ‘Requirements Document’ [8]. Although that document has already been published, SG38 continues to welcome all user input and suggestions, and to refine the specifications for the final product.

Once the requirements document was created, SG38 switched its focus to the design and implementation of the new nuclear database format. In order to expedite the work, the task was divided into several sub-tasks, each led by smaller groups from within SG38. These tasks include:

- Organizing the top-level hierarchy for storing evaluated nuclear reaction data.
- Designing basic data containers.
- Creating a database to store properties of particles, including nuclear excited states, masses, spins and decay information.
- Defining a common interface for infrastructure codes. The interface should support basic reading and writing of the new format, plotting, checking for physical accuracy, and processing into other forms for use by transport simulation codes. Where possible, existing infrastructure that handles ENDF data should be modified to also handle the new format and to implement the shared interface.
- Designing and implementing an API for reading and writing data in the new format. The API will eventually need to be implemented in multiple computer languages to support different users.
- Designing and building quality assurance tests to ensure that data files in the new format are correctly formatted, contain physically valid data, and can be used in application codes. Quality assurance includes physical tests for individual evaluations and a suite of tests such as critical assembly calculations.
- Establishing how the new format will be governed and ensuring proper documentation.

Generalized Nuclear Data (GND) and FUDGE

The work of SG38 has been influenced by the development at LLNL of a new hierarchical format for storing nuclear data, called Generalized Nuclear Data or GND [9]. Together with a code infrastructure called FUDGE (For Updating Data and Generating Evaluations), GND is being considered by SG38 as a possible starting point toward meeting all the requirements to serve as a replacement for ENDF.

GND files are inherently hierarchical, and can currently be stored in both XML and HDF5. The FUDGE platform can read and write GND files (in XML). Once data are read into FUDGE, they can be modified, plotted, checked for physical accuracy, processed into other formats, etc. FUDGE also includes the capability to translate ENDF data both to and from GND. This translation tool has already proved useful as a way to check the contents of ENDF evaluations.

Readers interested in seeing ENDF data translated into GND, or in interacting with the FUDGE infrastructure are encouraged to download FUDGE from https://ndclx4.bnl.gov/gf/project/gnd/. For more details about the capabilities of the FUDGE software system, please also see the contribution by B. Beck in these proceedings.

GND and FUDGE already contain many of the features described on the list of requirements assembled by SG38. However, the format specifications from SG38 are still in a state of flux while the subgroup determines the best model for a new format. As SG38 specifications are finalized, GND will be extended and modified as necessary to meet the final specifications. The authors expect that this modification can be performed within a reasonable time frame.

ENDNOTES

[a]GND actually defines a *structure* rather than a *format*. The structure can be stored in multiple different formats, including 'GND/XML' and 'GND/HDF5'.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

REFERENCES

1. A.TRKOV, M.HERMAN and D.BROWN (eds.), “ENDF-6 Formats Manual: Data Formats and Procedures for the Evaluated Nuclear Data Files ENDF/B-VI and ENDF/B-VII”, BNL-90365-2009 Rev.2, BNL (2012). (see http://www.nndc.bnl.gov/csewg/docs/endf-manual.pdf).
2. "Monte Carlo N-Particle Transport Code", https://mcnp.lanl.gov
3. WPEC Subgroup #38 (see http://www.oecd-nea.org/science/wpec/sg38/).
4. The HDF Group. Hierarchical Data Format version 5, 2000-2010. (see http://www.hdfgroup.org/HDF5)
5. "Experimental Nuclear Reaction Data", https://www-nds.iaea.org/exfor/exfor.htm
6. "Evaluated Nuclear Structure Data File", http://www.nndc.bnl.gov/ensdf/
7. "Reference Input Parameter Library", https://www-nds.iaea.org/RIPL-3/
8. "Requirements for a new nuclear data structure", produced by WPEC SG38 (2013). (see https://www.oecd-nea.org/science/wpec/sg38/requirements.pdf)
9. C.M. MATTOON et al., "Generalized Nuclear Data: A New Structure (with Supporting Infrastructure) for Handling Nuclear Data", *Nuclear Data Sheets* **113**, 3145 (2012).